

Einführung Datenmanagementplan

1. Workshop: „Das Konzept Datenmanagementplan“ 11.3.2021

18.03.21

Prof. Dr. Mirjam Blümm

Institut für Informationswissenschaft / Advanced Media Institute

Seite 1



Technology
Arts Sciences
TH Köln

Leitfragen

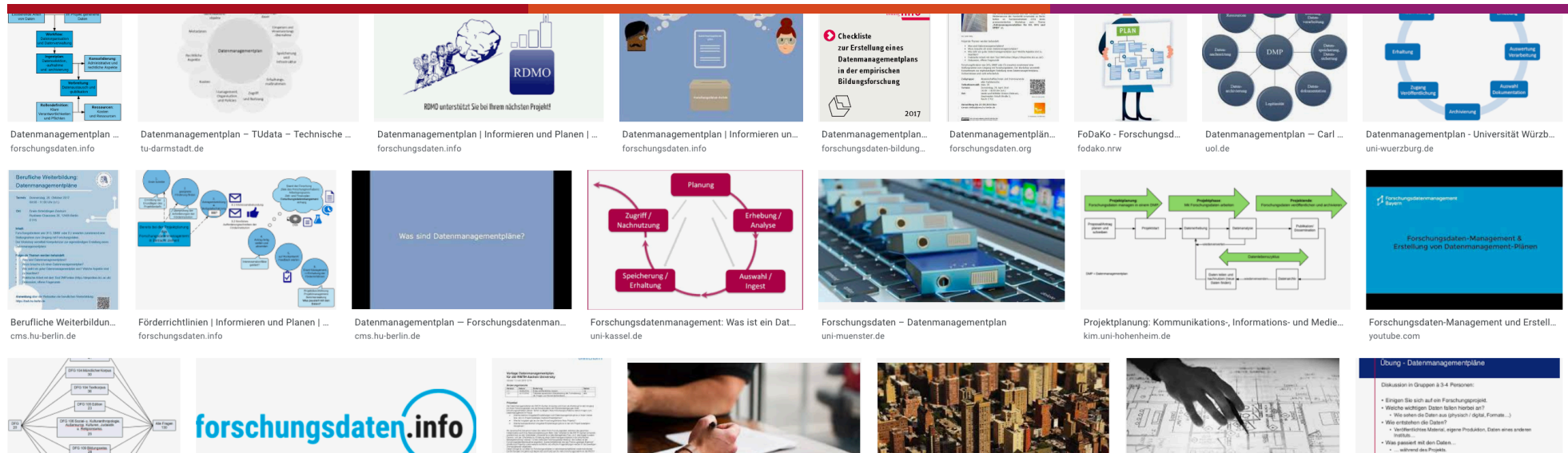
Was ist ein DMP?

Wozu dient ein DMP?

Warum wird die Forschung durch ein DMP besser?

Welche Bestandteile hat ein DMP?





(Bildquelle: Google Bildersuche „DMP“)

Was ist ein DMP?

Was sind „Daten“? | Was heißt „Management“? | Was ist der „Plan“?

18.03.21

Prof. Dr. Mirjam Blümm

Institut für Informationswissenschaft / Advanced Media Institute

Seite 3



**Technology
Arts Sciences
TH Köln**

Beispiele für (Forschungs-)Daten

- Messdaten
- Laborwerte
- Daten aus Umfragen
- Methodische Testverfahren (z.B. Fragebögen, Simulationen)
- Objekte (z.B. Sammlungen)
- Text
- Bilder, Digitalisate
- Audio- / Video-Daten
- Metadaten, Normdaten, Bibliographische Daten
- Programme, Software
- Datenbanken
- ...



Was sind (Forschungs-)Daten?

Forschungsdaten sind Daten, die im Zuge wissenschaftlicher Vorhaben z.B. durch Digitalisierung, Quellenforschungen, Experimente, Messungen, Erhebungen oder Befragungen entstehen.

Allianz der deutschen Wissenschaftsorganisationen
(<https://www.allianzinitiative.de/archiv/forschungsdaten/>)

Weitere Definitionen vgl. <https://www.forschungsdaten.info/themen/informieren-und-planen/was-sind-forschungsdaten/>



Was heißt Forschungsdatenmanagement?



Forschungsdatenmanagement (FDM) umfasst die Prozesse der Transformation, Selektion und Speicherung von Forschungsdaten mit dem gemeinsamen Ziel, diese langfristig und personenunabhängig zugänglich, nachnutzbar und nachprüfbar zu halten. Dazu können an allen Punkten des Datenlebenszyklus strukturierte Maßnahmen ergriffen werden, die geeignet sind die wissenschaftliche Aussagekraft von Forschungsdaten zu erhalten, deren Zugänglichkeit durch Dritte für Auswertung und Analyse zu bewahren und die Nachweiskette zu sichern.

<https://www.forschungsdaten.info/themen/planen-und-strukturieren/wie-beginnt-forschungsdatenmanagement/>

Bildquelle: <https://www.forschungsdaten.info/themen/informieren-und-planen/datenlebenszyklus/>

Was ist ein DMP?

- strukturierte Dokumentation wie Forschungsdaten erhoben, aufbereitet, zugänglich gemacht und archiviert werden
- umfasst alle Schritte des Datenlebenszyklus / Maßnahmen zur Zugänglichkeit und Erhaltung von Forschungsdaten
- bezieht sich i.d.R. auf die Daten eines (wissenschaftlichen) Projekts, Forschungsvorhabens
- legt fest, wie mit Daten **während und nach** der Projektlaufzeit umgegangen wird gemäß der projektspezifischen Anforderungen
- sollte möglichst **vor** der Projektlaufzeit erstellt werden
- muss **während** der Projektlaufzeit ergänzt und angepasst werden



Bildquelle: Dominik Schmitz, Daniela Hausen, Ute Trautwein-Bruns. Datenmanagement nach Plan. RWTH Aachen University. 2018. Verfügbar unter DOI: 10.18154/RWTH-2018-231100. (<https://youtu.be/XummkIOmMlg> YouTube, 11:34)

Ein Beispiel: Muster-Datenmanagementplan für einen DFG-Antrag

Kerstin Helbig:

Muster-Datenmanagementplan für
einen DFG-Antrag (Humboldt-
Universität zu Berlin, pdf)

<https://www.cms.hu-berlin.de/de/dl/dataman/muster-dmp-dfg/view>

Muster-Datenmanagementplan

Projektname: Umfrage zum Umgang mit Forschungsdaten an der HU

Projekt-ID/Förder-ID: XXX-YZ-2016

Primärforscher/Wissenschaftler: Kerstin Helbig

ID Primärforscher/Wissenschaftler: <http://orcid.org/0000-0002-2775-6751>

Kontakt: Tel. +49(0)30 2093-70072, Kerstin.Helbig@cms.hu-berlin.de

Projektbeschreibung: Das Projekt befasst sich mit dem Umgang mit Forschungsdaten von Wissenschaftlerinnen und Wissenschaftlern sowie Promovierenden, die an der Humboldt-Universität zu Berlin tätig sind. Via E-Mail und mittels Onlinefragebogen werden die Forscher zu ihrem aktuellen Umgang mit Forschungsdaten und eventuellen Beratungs- und Servicelücken befragt. Die Datensammlung dient der Feststellung des Bedarfs an Unterstützung und notwendiger zukünftiger Beratungs- und Dienstleistungen im Bereich Forschungsdatenmanagement.

Forschungsförderer: Deutsche Forschungsgemeinschaft

Förderprogramm: LIS-Förderprogramm „Informationsinfrastrukturen für Forschungsdaten“ - Forschungsdaten in der Praxis

Relevante Policies: Grundsätze zum Umgang mit Forschungsdaten an der Humboldt-Universität zu Berlin; Open-Access-Erklärung der Humboldt-Universität zu Berlin

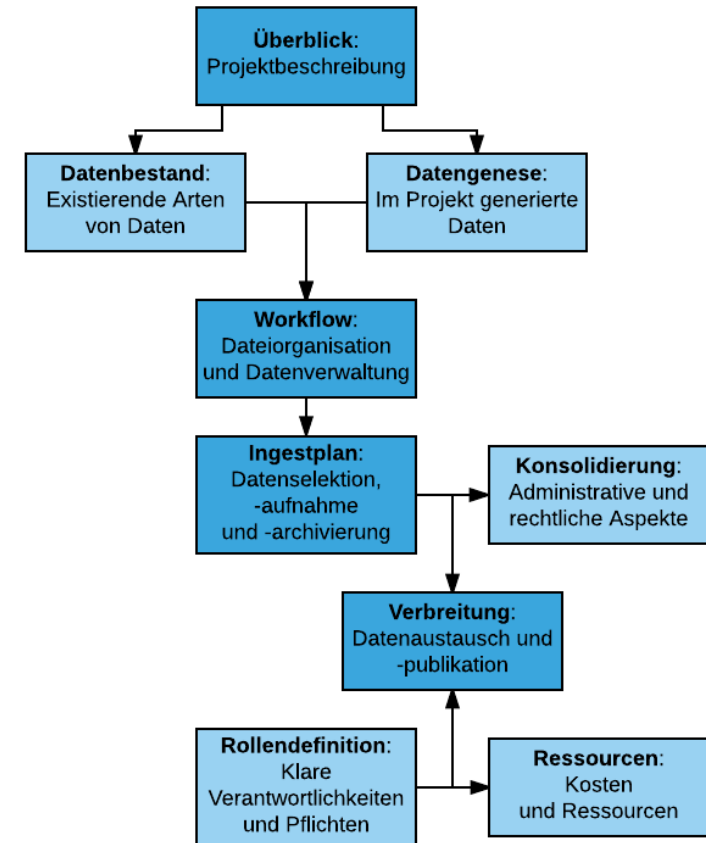
Beschreibung der zu erhebenden Forschungsdaten

Die Forschungsdaten werden über einen Onlinefragebogen erhoben. Hierzu wird die vom Computer- und Medienservice der Humboldt-Universität zu Berlin bereitgestellte Software LimeSurvey genutzt. Die Teilnehmer werden auf eine spätere Veröffentlichung unter



Wozu dient ein DMP?

- (Nach-)Nutzbarkeit von Daten (projektintern und darüber hinaus)
- Anforderung von Forschungsförderungsorganisationen: DMP als Teil des Antrags für bestimmte Förderlinien (DFG, EU, Volkswagenstiftung)
- Anforderung der guten wissenschaftlichen Praxis
- dauerhafte Aufbewahrung der verwendeten und erstellten Daten (min. 10 Jahre)



Bildquelle: <https://www.forschungsdaten.info/themen/planen-und-strukturieren/datenmanagementplan/>

Warum wird die Forschung durch einen DMP besser?



- klare Datenorganisation hilft den Überblick zu behalten (auch bei Personalwechsel)
- Speicher- und Archivierungsstrategie beugt Datenverlust vor
- klare Verantwortlichkeiten erleichtern die Durchführung des Projekts
- Sicherung der guten wiss. Praxis (u.a. Transparenz, Qualität)
- sorgt für wiss. Credits (Datenpublikation, Zitation von Forschungsdaten, ...)
- umfassende Dokumentation (incl. Reporting)

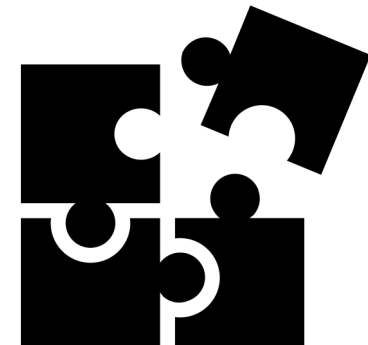


- Datenmanagement ist zeitaufwendig
- erfordert kontinuierlich Personalressourcen und ggf. Infrastruktur für die Erstellung und Pflege
- Vorlagen / Checklisten für DMP müssen ggf. angepasst werden
- unterschiedliche Vorgaben je nach Förderer, müssen erfüllt und ggf. in Einklang mit institutionellen Polycys gebracht werden

Vgl. auch
<https://www.forschungsdaten.info/theme/n/informieren-und-planen/datenmanagementplan/>

Welche Bestandteile hat ein DMP?

- Beschreibung der (zu erwartenden) Daten
- Rechtliche und ethische Aspekte (z.B. Datenschutz, Rechte an geistigem Eigentum, Einwilligung aller Beteiligten und Anerkennung der jeweiligen Leistung, ...)
- Bedingungen für die Nachnutzung (z.B. Lizenzen, ...)
- Metadatenpezifikation (z.B. verwendete Standards, ...)
- ggf. Beschreibung der verwendeten Methoden, Tools, Workflows, ...
- Planung der Langzeitarchivierung (z.B. Bitstream, Zugriff...)
- Regelung der Verantwortlichkeiten
- Administrative Angaben zum Projekt
- Kosten (z.B. für Langzeitarchivierung)
- ...



Kernelemente eines DMP (Empfehlungen Science Europe)



Science Europe: Practical Guide to the International Alignment of Research Data Management - Extended Edition (2021)

<https://scienceeurope.org/our-resources/practical-guide-to-the-international-alignment-of-research-data-management/>

Core Requirements for Data Management Plans (I)



1. Data description and collection or re-use of existing data

- a) How will new data be collected or produced and/or how will existing data be re-used?
- b) What data (for example the kinds, formats, and volumes) will be collected or produced?

2. Documentation and data quality

- a) What metadata and documentation (for example the methodology of data collection and way of organising data) will accompany data?
- b) What data quality control measures will be used?

3. Storage and backup during the research process

- a) How will data and metadata be stored and backed up during the research process?
- b) How will data security and protection of sensitive data be taken care of during the research?

Core Requirements for Data Management Plans (II)



4. Legal and ethical requirements, codes of conduct

- a) If personal data are processed, how will compliance with legislation on personal data and on data security be ensured?
- b) How will other legal issues, such as intellectual property rights and ownership, be managed? What legislation is applicable?
- c) How will possible ethical issues be taken into account, and codes of conduct followed?

Core Requirements for Data Management Plans (III)



5. Data sharing and long-term preservation

- a) How and when will data be shared? Are there possible restrictions to data sharing or embargo reasons?
- b) How will data for preservation be selected, and where will data be preserved long-term (for example a data repository or archive)?
- c) What methods or software tools will be needed to access and use the data?
- d) How will the application of a unique and persistent identifier (such as a Digital Object Identifier (DOI)) to each data set be ensured?

6. Data management responsibilities and resources

- a) Who (for example role, position, and institution) will be responsible for data management (i.e. the data steward)?
- b) What resources (for example financial and time) will be dedicated to data management and ensuring that data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?

Wann ist ein DMP ein guter DMP?

DMP Evaluation Rubric

DMP Question	DMP Guidance	Performance Levels	
GENERAL INFORMATION			
Guidance for Researchers		Sufficiently Addressed The DMP...	Insufficiently Addressed The DMP...
Administrative information	<ul style="list-style-type: none"> Provide information such as name of applicant, project number, funding programme, version of DMP. 	<ul style="list-style-type: none"> Contains the minimal information required to identify the applicant and the references of the project. 	<ul style="list-style-type: none"> Provides no or limited information, which makes it hard to identify who is responsible for the project.
1 DATA DESCRIPTION AND COLLECTION OR RE-USE OF EXISTING DATA			
Guidance for Researchers		Sufficiently Addressed The DMP...	Insufficiently Addressed The DMP...
1 a How will new data be collected or produced and/or how will existing data be re-used?	<ul style="list-style-type: none"> Explain which methodologies or software will be used if new data are collected or produced. State any constraints on re-use of existing data if there are any. Explain how data provenance will be documented. Briefly state the reasons if the re-use of any existing data sources has been considered but discarded. 	<ul style="list-style-type: none"> Gives clear details of where the existing data come from and how new data will be collected or produced. It clearly explains methods and software used. Explains, if existing data are re-used, how these data will be accessed and any constraints on their re-use. Explains clearly, if applicable, why new data must be collected, instead of re-using existing data. 	<ul style="list-style-type: none"> Provides little or no details on where the data come from and what data will be collected or re-used. Does not, if applicable, provide sufficient rationale for generating new data.

Science Europe: Practical Guide to the International Alignment of Research Data Management - Extended Edition (2021) ab S. 31
<https://scienceeurope.org/our-resources/practical-guide-to-the-international-alignment-of-research-data-management/>





Vielen Dank für Ihre Aufmerksamkeit.

Gibt es noch Fragen?

Prof. Dr. Mirjam Blümm

Technische Hochschule Köln

Fakultät für Informations- und Kommunikationswissenschaften – Fakultät für Informatik und Ingenieurwissenschaften

mirjam.bluemm@th-koeln.de

18.03.21

Prof. Dr. Mirjam Blümm

Institut für Informationswissenschaft / Advanced Media Institute

Seite 17

Technology
Arts Sciences
TH Köln

Quellen

- forschungsdaten.info
- [Jens Ludwig, Harry Enke: Was sind Forschungsdaten?](#) In: Lehrbuch Forschungsdatenmanagement (2016) https://handbuch.tib.eu/w/Lehrbuch_Forschungsdatenmanagement
- Dominik Schmitz, Daniela Hausen, Ute Trautwein-Bruns. Datenmanagement nach Plan. RWTH Aachen University. 2018. Verfügbar unter DOI: 10.18154/RWTH-2018-231100. (<https://youtu.be/XummkIOmMlg> YouTube, 11:34, deutsch)
- Kerstin Helbig: Muster-Datenmanagementplan für einen DFG-Antrag (Humboldt-Universität zu Berlin, pdf) <https://www.cms.hu-berlin.de/de/dl/dataman/muster-dmp-dfg/view>
- Kerstin Helbig; Katja Krause; Carolin Kruse; Florian Rehak; Gianpiero Tari (HU Berlin). [Was sind Datenmanagementpläne?](#) Verfügbar unter: <https://doi.org/10.18450/dataman/91> (4:11, deutsch)
- Science Europe: Practical Guide to the International Alignment of Research Data Management - Extended Edition (2021) <https://scienceeurope.org/our-resources/practical-guide-to-the-international-alignment-of-research-data-management/>

